

This is a repository copy of *How certain are we about the certainty-equivalent long term social discount rate?*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/112082/>

Version: Accepted Version

Article:

Freeman, Mark Charles orcid.org/0000-0003-4521-2720 and Groom, B (2016) How certain are we about the certainty-equivalent long term social discount rate? *Journal of Environmental Economics and Management*. pp. 152-168. ISSN 0095-0696

<https://doi.org/10.1016/j.jeem.2016.06.004>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

How certain are we about the certainty-equivalent long term social discount rate?

Mark C. Freeman and Ben Groom¹

Abstract

Theoretical arguments for using a term structure of social discount rates (SDR) that declines with the time horizon have influenced Government guidelines in the US and Europe. The certainty equivalent discount rate that often underpins this guidance embodies uncertainty in the primitives of the SDR, such as growth. For distant time horizons the probability distributions of these primitives are ambiguous and the certainty equivalent itself is uncertain. Yet, if a limited set of characteristics of the unknown probability distributions can be agreed upon, ‘sharp’ upper and lower bounds can be defined for the certainty-equivalent SDR. Unfortunately, even with considerable agreement on these features, these bounds are widely spread for horizons beyond 75 years. So while estimates of the present value of intergenerational impacts, including the social cost of carbon, can be bounded in the presence of this ambiguity, they typically remain so imprecise as to provide little practical guidance.

Key Words Declining discount rates, Distribution uncertainty, Social Cost of Carbon.

JEL classification H43, Q51.

1 Introduction

The outcome of cost-benefit analysis of public projects with intergenerational consequences is notoriously sensitive to the social discount rate (SDR) employed. Small variations in

¹Mark C. Freeman (Corresponding Author), School of Business and Economics, Loughborough University, Leicestershire LE11 3TU, United Kingdom. Email: M.C.Freeman@lboro.ac.uk. Ben Groom, School of Geography and Environment, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom. Email: b.groom@lse.ac.uk.

assumptions about the appropriate SDR can therefore lead to very different policy recommendations for the preservation of natural resources and environmental quality, including the retention of biodiversity (Freeman & Groom (2013)) and the case for mitigating against greenhouse gas emissions (e.g. Nordhaus (2007), Stern (2008)).

This policy-sensitivity is particularly problematic because the primitives that underlie the long-term discount rate are difficult to determine. For example, the growth rate of aggregate consumption and the rate of return to capital over the next four centuries are essentially unknown today, since they depend on a number of unpredictable events including technological advances, political and social unrest, environmental change and even pandemics (e.g. Almond (2006)).

A typical way to approach long-term discounting is to calculate a ‘certainty equivalent’ social discount rate, a single rate which embodies uncertainty in the SDR primitives. Yet even though uncertainty is taken into account, such calculations assume a fanciful level of predictive power, since they assume perfect knowledge of the relevant probability distributions. In the context of intergenerational decision-making, the probabilities associated with different future states of the world are thought to be ambiguous at best, and at worst unknown.² Consequently, the certainty equivalent discount rate is itself uncertain.

In this paper we make a contribution to the literature on social discounting under uncertainty by calculating empirical ‘sharp’ upper and lower bounds for the certainty-equivalent social discount rate when we have imperfect knowledge of probability distributions of SDR primitives. Such bounds can be calculated if decision-makers are willing to assume partial, but not complete, agreement on some characteristics of these distributions. The existence of sharp bounds is the good news. The bad news is that these bounds are typically very wide and fail to provide precise calculations of present values.

These findings are important because the burgeoning literature on the term structure of

²We use the term ‘ambiguous’ in this paper in the sense of (Klibanoff et al. 2011, p.400) “that this definition is characterized by, roughly, disagreement in the probability assigned to an event by the various probability measures that are subjectively relevant.”

social discount rates, expertly reviewed by Gollier (2012) and Arrow et al. (2014), has been highly influential at a policy level. The message coming from these contributions is that, for risk free projects, the term structure should be declining with the time horizon. This view is exemplified by a recent *Policy Forum* article in *Science*, in which it is argued that where we are uncertain about the future “there are compelling arguments for using a declining discount rate schedule” (Arrow et al. 2013, p. 350). As a consequence of these theoretical advances, declining discount rates (DDRs) can now be found in government guidelines in the UK and France, influence recommendations in the US (Cropper et al. (2014)), and lie behind recent advice given to the Norwegian, Danish and Dutch governments. In the UK, DDRs have been used in the governmental economic analysis of the High Speed 2 (HS2) rail link and for capital budgeting purposes by the Nuclear Decommissioning Authority. DDRs have already had policy impact.

The DDRs that appear in government guidelines are typically based on certainty-equivalent discount rates which reflect uncertainty in the future or disagreement among experts on the appropriate discount rate, perhaps for ethical reasons. An influential set of arguments supposes that for some x , the different definitions of which are reviewed in subsequent sections, the present value, p_H , of a certain \$1 arriving at time H is given by $p_H = E[\exp(-Hx)]$. The H -period certainty-equivalent discount rate, R_H , is then defined through the relationship $\exp(-HR_H) = p_H$. Exponential functions are convex, and so, by Jensen’s inequality, $E[\exp(-Hx)] \geq \exp(-HE[x])$: uncertainty over x raises the present value, p_H , and lowers the discount rate, R_H . The magnitude of this effect becomes greater the more uncertain we are about x , and the more convex the exponential function, the latter being determined by the parameter value H . As a consequence R_H declines with the time horizon until in the limit, as $H \rightarrow \infty$, it approaches the lowest possible outcome for x .³

³A numerical example illustrates the mechanics of the result. Suppose that, with equal probability, x will either take the value of 2% or 6%. The social value of \$1 delivered at time H with certainty is then given by the expected present value under these two outcomes, $p_H = 0.5(\exp(-0.02H) + \exp(-0.06H))$, resulting in $R_1 = 3.98\%$, $R_{50} = 3.13\%$, $R_{100} = 2.67\%$ and $R_{400} = 2.17\%$. The $x = 6\%$ outcome is, through the power of exponential discounting, given increasingly less voice in the social valuation p_H as H gets larger. For horizons of a century or more, to good approximation, its contribution to p_H becomes so small that it can

While this may seem like a narrowly-defined structure, it has several different interpretations depending on the approach taken to social discounting and DDRs. Its most famous use stems from Weitzman’s (2001) ‘Gamma Discounting’ paper. Here, x was interpreted as reflecting different expert opinions on the value of the discount rate itself. In this context, the justification for using the formula $p_H = E[\exp(-Hx)]$ remains controversial. This is partly because its connection with utility theory was not made clear at the time, and partly that more recent theoretical motivations rely on quite restrictive assumptions. We discuss this point in detail in Sections 2 and 3. Less well known is that x can also be interpreted through the Social Rate of Time Preference (SRTP) in a more standard consumption-based Ramsey asset pricing framework. A proof of this proposition is given in Section 3. A third interpretation of x , also discussed in Section 3, is that it represents the average return to risk-free capital over the horizon of the cash flow. Since they all have a similar expectations structure, the methods that we describe for deriving the sharp bounds for the certainty equivalent discount rate can be equally well applied to any of these interpretations.

Putting any of these interpretations into practice requires assumptions about the uncertainty surrounding the primitives of the SDR that are contained in x , through its probability density function (pdf), $f(x)$. The main approach taken so far is to parameterize $f(x)$ and treat this distribution as if it is perfectly identified. Yet, because our knowledge of the future is nowhere near as precise as this approach would suggest, a more realistic starting point would be to admit that we do not, perhaps cannot, know the true nature of $f(x)$ over time horizons of many decades or centuries. For very long-term decisions the context is one of uncertainty and ambiguity. We are not alone in thinking this. Pindyck (2015) recently made a similar point in relation to social discounting, while Iverson (2013) and Traeger (2014) both take ambiguity as their starting points. Yet the approach that we take assumes far more knowledge about the future than Knight (1921), who would maintain that true uncertainty is immeasurable. To reduce the problem we imagine a situation where the social planner

be ignored altogether with $p_H \approx 0.5 \exp(-0.02H)$ and $R_H \approx 0.02 - \ln(0.5)/H$. In the limit, as $H \rightarrow \infty$, $R_H \rightarrow 2\%$.

gathers a panel of economists who, while accepting that it will be impossible to agree on the precise distribution of x , are nonetheless tasked with identifying the *set* of density functions such that all members agree that $f(x)$ is a member of this set.⁴ Agreement in this context takes the form of agreeing characteristics shared by all distributions within a set. For instance, the set may contain probability distributions which share the same first K moments, or alternatively the same values for particular quantiles.⁵ They might, instead, be members of a family of distributions, such as the gamma or Wald (Inverse Gaussian) distribution. Of course, the number of pdfs which share the basic information that defines the set is infinite, and so there is significant room left for disagreement. Yet, as we show, given some level of agreement on shared characteristics, the social planner can determine the highest and lowest certainty equivalent social discount rate that is consistent with the agreed set of pdfs, for each horizon H . In contrast to Iverson (2013) and Traeger (2014), therefore, our focus is on proving a range of potential values for R_H which the social planner cannot easily dismiss, rather than determining a single precise schedule of discount rates for an ambiguity averse decision maker.

Our central conclusion is that, despite limiting the ambiguity over the future by requiring significant agreement on some characteristics of $f(x)$, the range of values that R_H might reasonably take remains extremely wide for horizons of three-quarters of a century or more irrespective of the theoretical interpretation of x . This result is primarily driven by uncertainty over the properties of the extreme left-hand tail of $f(x)$ even when there is widespread consensus on other characteristics of the distribution. While clearly distinct, this finding echoes the substantial existing literature explaining the key role played by extreme outcomes in the economics of climate change.⁶

⁴While it is possible that these assumptions will be falsified with the benefit of hindsight, we assume that the social planner is willing to make decisions on the basis of assumptions about $f(x)$ that are sufficiently uncontroversial for reasonable people to be able to agree upon them today. O'Hagan et al. (2006) provides a detailed review of how experts' probability judgements might be assessed for this purpose.

⁵To avoid issues around infinities, as famously discussed in a related context through the 'dismal theorem' of Weitzman (2009), we assume throughout that the first K moments of $f(x)$ are finite and that, more generally, its moment generating function is defined.

⁶See, for example, papers published by Professors Weitzman, Nordhaus and Pindyck summarizing a

Three policy applications are then presented in which the present values of long-term cash flow schedules are calculated: (i) the damages caused by carbon emissions; (ii) the estimated benefits of Phase 1 of the HS2 rail link; and (iii) the costs of decommissioning the previous generation of nuclear power stations within the UK. As is to be expected, the less we agree upon about the future, the wider are the bounds of the certainty equivalent discount rate and the more uncertain we are about the ‘true’ present value. For example, using gamma discounting (Weitzman (2001)) as our underlying DDR model, even if there is agreement on the first four moments of $f(x)$, then the social cost of carbon (SCC) can lie anywhere within the interval \$13.6 per ton of carbon (/tC) and \$46.1/tC. This is a sobering result when one considers that even agreement on the first moment is likely to be an optimistic assumption. In this latter position of relative ignorance, the SCC can be anywhere between \$5/tC and \$190/tC.

The obvious conclusion for policy, therefore, is that social planners should be extremely cautious when making decisions on intergenerational matters based on calculating Net Present Values. Even if we make the bold claim that we know some minimal summary statistics of the uncertainty surrounding future growth or interest rates, and ignore broader issues such as true Knightian uncertainty, we can still only be sure that the appropriate discount rate will lie within very wide bounds. In such cases spot estimates of, say, the SCC will give a false impression of precision. So, while on the up-side our method provides a concrete way in which to set the range for sensitivity analysis in cost-benefit analysis, the broader message is that additional decision-making criteria are likely to be required for long-term projects.

2 DDRs under gamma discounting

Consider a social planner who has decided to implement a schedule of declining discount rates for calculating the present value of the certainty-equivalent cash flows that will arise from

symposium on “Fat tails and the economics of climate change” published in the *Review of Environmental Economics and Policy*, 5(2), 2011.

intergenerational investment projects. As discussed in the introduction such a policy can be motivated by an H -period certainty equivalent discount rate with the following structure:

$$R_H = -H^{-1} \ln (E [\exp (-Hx)]) \quad (1)$$

for some x whose probability or frequency distribution, $f(x)$, might be H -dependent. While the next section will go into detail about the theoretical justifications for this structure and the different associated interpretations of x , at this point it is helpful to briefly mention the main interpretations of x that appear in the literature. In one strand, x is interpreted as the average long-term risk-free rate of interest, for another it is the Social Rate of Time Preference (S RTP) from the standard Ramsey model, $S RTP = \rho + \gamma g$, where ρ is the pure rate of time preference, γ is the elasticity of marginal utility, and g is the real growth of per capita consumption. Given these different possibilities, the data on x could be elicited from survey responses, as in Weitzman (2001) and Drupp et al. (2015). Alternatively it might be derived from revealed preference, such as in Groom & Maddison (2013) for γ . The data could reflect normative views for ρ and γ or could represent forecasts of market interest rates and growth, perhaps based on historical data. Consequently, the expectations operator, $E[\cdot]$, can act either (as usual) as a probability weighted aggregator of the possible future realizations of a stochastic random variable, or as a weighted average across different opinions or characteristics. In the broader literature these distinct sources of data determine the way in which the expectation in equation 1 is calculated (Freeman & Groom (2015)).

In this section we view equation 1 through the lens of the ‘gamma discounting’ framework of Weitzman (2001). Here, x represents different expert opinions on the discount rate and $f(x)$ is the frequency distribution of those opinions. This framework is selected to explain our main point not because we believe that it is necessarily the best for determining the schedule of social discount rates, but instead because of its simplicity. This choice allows us to make our central point with minimal extraneous complexity before we extend the analysis

to other interpretations of equation 1 and x in Sections 3 and 4.

Gamma discounting used a survey of $N = 2160$ PhD level economists who were asked the question ‘*Taking all relevant considerations into account, what real interest rate do you think should be used to discount over time the (expected) benefits and (expected) costs of projects being proposed to mitigate the possible effects of global climate change?*’. Weitzman (2001) then argued that the social value, p_H , of \$1 that will arrive at time H with certainty should be determined by an average of the individual discount factors. If the response of expert i is denoted by x_i , the SDR is given by:

$$R_H = -\frac{1}{H} \ln \left(\sum_{i=1}^N w_i \exp(-Hx_i) \right) \quad (2)$$

where w_i is the weight placed on expert i by the social planner.⁷

Rather than applying this framework directly, Professor Weitzman instead noticed that the sample frequency distribution of responses was approximately gamma distributed, $\Gamma(\alpha, \beta)$, for rate parameter α and shape parameter β . Under this approximation, if equal weight is given to each expert, $w_i = 1/N$ for all i , equation 2 becomes:

$$R_H = -\frac{1}{H} \ln \left(\int_0^\infty e^{-Hx} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \right) = \frac{\alpha}{H} \ln \left(1 + \frac{H}{\beta} \right) \quad (3)$$

This gives a convenient closed-form solution for the term structure of the social discount rate that depends only on the two parameters of the gamma distribution. For the purposes of this paper, ‘gamma discounting’ is defined as determining the SDR through equation 3.

Now assume that the social planner has decided to use equation 2 to determine the SDR, and then chosen to calibrate this equation using Weitzman’s survey data.⁸ However, she has

⁷A similar approach can also be applied when a social planner is constructing a global discount rate that reflects the potentially conflicting positions of different social groups or geographical regions. Similar to the case of expert opinion, $f(x)$ represents the weighted population frequency distribution of the different discount rates that would be applied by either a specific sub-group of society or international governments. See, for example, (Gollier 2012, Chapter 9).

⁸We delay discussion of the theoretical arguments for and against using equation 2 to the next section. Choosing to calibrate with Weitzman’s data is also controversial. Dasgupta (2008), for example, argues

serious reservations about applying equation 3 because she remains unconvinced that setting $w_i = 1/N$ is a natural choice. Her views here are influenced by Gollier & Zeckhauser (2005), Jouini et al. (2010), Heal & Millner (2014), Jouini & Napp (2014) and Freeman & Groom (2015), all of whom argue that experts should only be weighted equally under very restrictive assumptions — we will return to this issue in the next section. To seek further guidance, she brings together a panel of J economists to reach consensus on which properties of w_i , and hence $f(x)$ are sufficiently uncontroversial so as to provide a solid ground for policy decisions. The panel quickly agrees that all experts should be given non-negative weight, $w_i \geq 0$ for all i , and that the weights should sum to one. These choices ensure that equation 2 can be interpreted within the context of equation 1 when, for mathematical convenience, each x_i is treated as if it were a possible realization of a random variable that has probability w_i of occurring. In this context $f(x)$ has the interpretation of being a weighted sample frequency distribution.

Unfortunately, being economists, panel members fail to agree on the choice of weightings, w_i , so each of them proposes their own weighted distribution, $f_j(x)$ for $j \in [1, J]$, to the social planner. Disagreement in this case could reflect variations in the way that each panel member perceives the professional standing of each expert. Alternatively it might be that panel members have different preferences for the specific approaches to social discounting that are advocated by different groups of the sampled population.⁹ Finally, the proposed $f_j(x)$ could reflect the panel member’s position on how independent the expert opinions are from one another (Freeman & Groom (2015)). Deciding on these weights is ultimately a subjective professional choice, and there is no strong reason for the social planner to believe that the values selected by one panel member are in any sense better than those of another.

Therefore, in a spirit of open-mindedness, she views all panel members’ weighting choices

that a more democratic approach should be taken to social discounting than just canvassing the opinions of ‘experts’. Furthermore, a recent survey by Drupp et al. (2015) reports a very different spread of expert views on the SDR to Weitzman (2001).

⁹Drupp et al. (2015) provides a lengthy discussion of the underlying reasons why experts give conflicting advice on R_H , and presents detailed evidence concerning the spread of opinions on each of these matters.

as equally plausible, and will not look to further aggregate these conflicting opinions into a single discount rate schedule.

In order to provide useful input to policy further consensus is needed to stop cases of extreme disagreement.¹⁰ Specifically, the social planner asks that weights are chosen to ensure that certain statistical properties of $f_j(x)$ are the same across all panel members, j . While not exhaustive, these might relate to agreeing on some combination of: (i) the support of $f(x)$, $[a, b]$, which corresponds to the lowest and highest expert opinions that can be given non-zero weight; (ii) the moments of $f(x)$; (iii) the quantiles of $f(x)$; and/or (iv) whether a specific family of distributions can be used to describe $f(x)$. In this set-up, our central question becomes: given agreement between panel members on some, but not all, of the properties of $f(x)$, what are the maximum and minimum values that R_H can take that are consistent with this level of agreement? Following on from this, has the panel reached sufficient consensus so that the remaining disagreement over R_H does not ‘matter’ from a policy perspective? These questions are answered in the following subsections.

2.1 Agreement on the moments of x

Initially, the panel are willing to accept that some features of the spread of weighted expert opinions on the discount rate, $f(x)$, can be derived from the $\Gamma(\alpha, \beta)$ distribution that approximates the equally-weighted sample frequency distribution of Weitzman’s survey data. The panel also agree that parameter values, $\alpha = 1.90$ and $\beta = 47.23$, can be estimated using a methods of moments approach. From this framework, agreement is reached that $a = 0\%$, the lower bound of the gamma distribution, and $b = 19.11\%$, which is the 99.9th percentile of the same distribution. While the raw survey data of Weitzman have responses below a and above b , the panel reaches agreement that the extreme outliers lie sufficiently far from the mean as to be given zero weight. Our results are much more sensitive to the choice of the former than the latter as $R_H \rightarrow a$ as $H \rightarrow \infty$. We return to this issue in Subsection 2.3.

¹⁰Such as one panel member placing all weight on the expert with the highest response and another placing all weight on the one with the lowest response.

The panel is also willing to estimate some or all of the mean, standard deviation, skewness, and excess kurtosis of $f(x)$ from this distribution. These values are respectively 4.03%, 2.92%, 1.45 and 3.15. However, this is the limit of the agreement that the panel members can reach.

Suppose first that the panel can only reach agreement that the mean of $x = \sum_i w_i x_i$ equals 4.03%, together with a and b . Then, at one extreme, one panel member might wish to place all weight on the expert whose opinion matches the agreed mean value of $\Gamma(\alpha, \beta)$, 4.03%.¹¹ This results in $R_H = 4.03\%$ for all H . At the other extreme, another panel member might think that the two most informed experts lie at the opposite far ends of the support, $a = 0\%$ and $b = 19.11\%$ and chooses weights on each expert so as to match the agreed mean. These weights turn out to be 78.9% and 21.1% respectively. In this case, the social value is the weighted average of the expert discount factors; $p_H = 0.789 + 0.211 \exp(-0.1911H)$, resulting in $R_{50} = 0.5\%$ and $R_{100} = 0.2\%$. Taking together the views of the two panel members, for cash flows one century from now, the social planner cannot discard the possibility that the appropriate discount rate might be as low as 0.2% or as high as 4%.

Given this wide range of possible values for R_H , the social planner asks the panel whether they might also be willing to accept the estimate of the standard deviation of x from $\Gamma(\alpha, \beta)$. If so, the following two extreme distributions each meet this new consensus. For the first, the weighting is 96.4% at $x = 3.47\%$, with the remaining 3.6% at b . For the second, the weighting is 34.4% at a and 65.6% at 6.15%. Now R_{100} is equal to 3.5% for the first distribution and 1.1% for the second.

Again, the social planner feels that this does not provide a sufficiently narrow range to make important policy choices, so now asks the panel to agree on the skewness as well. In this case, R_{100} can lie in the range 1.3% to 2.7%. With kurtosis added, the range narrows

¹¹While Weitzman (2001) reports integer values for x_i , we assume that these round a continuum of responses over $[a, b]$. Placing all weight on one response, the median (not the mean), has been proved to be the optimal choice of SDR for a social planner who wishes to choose a democratic voting policy over optimal consumption paths (Millner & Heal (2014)).

further to 1.7% to 2.5%.¹²

The examples we have described of the extremes of disagreement that can persist over the weights may seem arbitrary, but they are not. As shown by Karlin & Studden (1966), in each case they represent the distributions which, while maintaining agreement on both the support and $K \in [1, 4]$ moments of $f(x)$, give the absolute upper and lower ‘sharp’ bounds of R_H .¹³ A detailed discussion of this method is provided in Appendix A. The top graph in Figure 1 shows these sharp bound values of R_H for $H \in [1, 400]$ and $K \in [1, 4]$:

[Insert Figure 1 around here]

For $K \leq 2$, there is a wide potential range of R_H for all but the shortest horizons. For higher values of K , which correspond to greater levels of agreement amongst the panel, the bounds remain relatively close together for approximately the first 75 years. After that, however, the differences become clear, even for $K = 4$.

Our central point follows from Figure 1. The social planner has extracted very high levels of agreement out of the panel of economists advising her, more than is perhaps realistic. The panel must estimate the social discount rate using equation 2, agree that the highest and lowest possible values of x are 0% and 19.11% respectively, and then further agree on several moments of $f(x)$. Yet, despite this fanciful level of consensus, the sharp bounds for R_H remain widely spread, particularly for long time horizons.

2.2 Agreement on the quantiles

Given the top graph in Figure 1, the social planner is dissatisfied with the levels of uncertainty that remain over R_H when discussions focus around the moments of the distribution.

¹²For skewness, the upper bound is given when the probability density function has mass of 79.3% and 20.7% at $x = 2.54\%$ and 9.76% respectively. The lower bound is given when the probability density function has mass of 25.8%, 72.7% and 1.54% at $x = a$, 5.14% and b . When kurtosis is included the weights are 66.7%, 32.5% and 0.8% at $x = 2.08\%$, 7.68% and b for the upper bound and 17.6%, 74.4% and 8.0% for $x = a$, 4.08% and 12.45% for the lower bound.

¹³This relies on the observation that the set of functions $\{1, x, \dots, x^K, (-1)^{K+1} \exp(-Hx)\}$ for $H > 0$ and positive integer K is a Tchebycheff system.

She therefore returns to the panel of economists and asks instead that they consider the quantiles of $f(x)$, as well as its supports, $a = 0\%$ and $b = 19.11\%$. This decision might be motivated by the observation in (O'Hagan 2012, p.37) that “The quantitative judgements that are used in elicitation are almost invariably evaluations of probabilities. Although we might ask the expert to assess quantities such as means and variances, the evidence from the experimental literature is generally that these are evaluated less accurately than probabilities... Moments are cognitively more complex constructs and highly sensitive to the thickness of a distribution's tails”.

In the first instance, she asks that the panel agrees on the terciles ($Q = 3$) of the distribution. With the panel again basing their estimates of the quantiles on $\Gamma(\alpha, \beta)$ estimated using Weitzman's (2001) sample data, consensus is reached that weighting summing to $1/3$ should be applied to expert opinions that lie within each of the following three ranges: $x < 2.35\%$, $2.35\% < x < 4.61\%$, and $x > 4.61\%$.

The sharp lower and upper bounds for R_H contingent on this level of agreement are now straightforward to derive. They respectively correspond to cases when all permissible weight is placed at the lower and upper bounds of each quantile range. In the example to hand, this is one-third weightings at $a = 0\%$, 2.35% and 4.61% for the lower bound and the same weighting at 2.35% , 4.61% and $b = 19.11\%$ for the upper bound. The social value is given by the weighted average discount factor, so the bounds for p_H at $H = 100$ years are:

$$\frac{1 + \exp(-2.35) + \exp(-4.61)}{3} > p_{100} > \frac{\exp(-2.35) + \exp(-4.61) + \exp(-19.11)}{3} \quad (4)$$

resulting in $1.00\% < R_{100} < 3.35\%$.

The social planner again asks the panel to increase its level of agreement by raising the value of Q to first five and then seven. In the former case, the top and bottom deciles are included (1.02% and 7.93% for $\Gamma(\alpha, \beta)$). For the latter, the first and 99^{th} percentiles are also captured (0.27% and 13.67% for $\Gamma(\alpha, \beta)$).¹⁴ When $Q = 5$, $R_{100} \in [1.52\%, 2.79\%]$, while

¹⁴These quantile levels are chosen to match the Intergovernmental Panel on Climate Change Fifth As-

when $Q = 7$, $R_{100} \in [1.63\%, 2.72\%]$. The bottom graph in Figure 1 plots the strict upper and lower bounds for R_H for $Q \in \{3, 5, 7\}$ and $H \in [1, 400]$.

In contrast to the top graph of Figure 1, for short time horizons, the bounds remain clearly distinct. This is because agreement on quantiles does not necessarily lead to agreement on the moments of the distribution. The individual $f(x)$ that create the bounds have different mean values. Since $R_H \rightarrow E[x]$ as $H \rightarrow 0$, this leads to wide bounds for short horizons. Crucially though, the term structures remain widely separated for all horizons. This makes it clear that while focussing on quantiles may be cognitively easier for the members of the expert panel than focussing on the moments, it provides no decision-making advantage from the perspective of the social planner.

2.3 Families of $f(x)$

In the analysis described so far, the sharp bounds for R_H are derived from distributions that place non-zero weight on only a very limited number of experts. The social planner may, though, agree with Weitzman (2001) that all experts should be heard in the social discount rate, and that setting $w_i = 1/N$ for all i is a sensible choice. However she still seeks guidance from the panel on how $f(x)$ should be estimated in this case.

One subset of the panel is willing to accept that $f(x)$ is best described by a gamma distribution, but would prefer that α and β are estimated through a maximum likelihood method. Based on the strictly positive responses to Weitzman’s survey, this gives $\alpha = 2.54$ and $\beta = 63.08$. The implied mean and standard deviation of this distribution are 4.03% and 2.53% respectively, giving a lower spread than when the parameter values are estimated by method of moments.

The second subset of the panel argue instead that Weitzman’s data is better fitted by a Wald (Inverse Gaussian) distribution than a gamma distribution. This is also supported on the interval $[0, \infty)$ and is positively skewed, and a Kolmogorov-Smirnov test weakly indicates

assessment Report’s (IPCC AR5) probabilistic definitions of the terms “Virtually certain”, “Very likely”, “Likely”, “About as likely as not”, “Unlikely”, “Very unlikely” and “Exceptionally unlikely”.

that it fits Weitzman's survey data better than a gamma distribution. Based on a method of moments approach, the distribution is calibrated with shape parameter $\lambda = 0.0767$ and mean $\phi = 0.0403$. Analogous to equation 3, from the moment generating function of this distribution:

$$R_H = -\frac{1}{H} \ln \left(\int_0^\infty e^{-Hx} \sqrt{\frac{\lambda}{2\pi x^3}} \exp \left(-\frac{\lambda (x - \phi)^2}{2\phi^2 x} \right) dx \right) = \frac{\lambda}{H\phi} \left(\sqrt{1 + \frac{2\phi^2 H}{\lambda}} - 1 \right) \quad (5)$$

Again, the term structure of discount rates can be simply derived in closed form based on the two parameter values of the underlying distribution.

A third subgroup of the panel disagrees with both of these choices of distribution for $f(x)$ since they are supported on the positive real line. Selecting these distributions effectively places zero weight on the three respondents to Weitzman's survey who gave negative values, the lowest of which was -3%, as well as the additional 46 who expressed their opinion that the social discount rate should be zero. This subgroup of the panel argues that more attention should be paid to the lower bound, a , particularly as Weitzman (1998) proves that $\lim_{H \rightarrow \infty} R_H = a$. They believe that $f(x)$ is best approximated by the raw sample frequency distribution of responses and that equation 2 should be applied directly to the data with $w_i = 1/N$ for all i . The top graph in Figure 2 displays the cumulative distribution functions for these different choices, and from this it is clear that they are highly similar and there is no obvious a priori reason why the social planner should have a strong preference for one over the others.

[Insert Figure 2 around here]

The bottom graph in this figure presents the term structure of discount rates in each case. For the gamma and Wald distributions, the term structures are closer together than those shown in Figure 1. However, for long time horizons they are clearly distinct. Using the sample frequency distribution leads to a very different term structure to any of the others reported in this paper. Beyond 250 years, R_H becomes negative, which implies

that the social planner should compound the cash flows received after this horizon.¹⁵ This illustrates the sensitivity of the analysis to the lower support, a , for $f(x)$, since this is the only distribution that is considered in this section that allows x to take negative values.

While clearly distinct, this analysis draws parallels with the literature that discusses the role played by extreme events in the economics of climate change. In particular, when explaining his ‘dismal theorem’, (Weitzman 2011, p.287) argues that if the probability density function of the natural logarithm of disutility is upper fat-tailed, then “the willingness to pay (WTP) to avoid extreme climate changes (is) very large, indeed arbitrarily large if the coefficient of relative risk aversion is bounded above one.” In such cases it is vital to understand the possible extremes of climate change damage, the worst case scenarios for consumption, and the characteristics of utility in the event of an environmental disaster in order to best inform policy. In our framework, there is no uncertainty over the damages as the analysis is being undertaken on a certainty-equivalent basis, and, in this section, no explicit consideration of consumption and utility. However, the wide bounds for R_H that we report at long horizons are largely driven by the different weights placed on the experts who gave the lowest value responses to the gamma discounting survey. Forcing agreement on the first four moments, seven quantiles, or broad families of distribution does not sufficiently constrain these weights to precisely define the certainty equivalent discount rate. It is uncertainty over the extreme left-hand tail of $f(x)$ that is primarily driving the results in this paper.

2.4 Impact on Net Present Values

Having determined a range of possible paths for declining discount rates based on gamma discounting, we now turn to how sensitive intergenerational valuations are to the schedule chosen. Table 1 reports the present value of \$1 million that is to be received with certainty

¹⁵From equation 2, $R_H < -H^{-1} \ln(\exp(0.03H)/2160)$, where the right-hand side represents the contribution of the single expert who gave a response of -3% for the social discount rate. The right-hand side is negative if $H > \ln(2160)/0.03 = 256$. This corresponds closely to the point where R_H crosses the x -axis in the bottom graph of Figure 2. One expert alone, therefore, explains a key characteristic of this figure.

in either 100, 200 or 400 years under each different scenario.

[Insert Table 1 around here]

Panel A presents each valuation under the baseline model. Panels B and C show that, as is to be expected, the range that p_H can take becomes more widely spread the longer the time horizon, H , and the less the level of agreement, K or Q , that is reached by the panel of economists. Even at 100 years, and with the greatest levels of agreement considered, there is more than a twofold difference between the highest and lowest valuation whether discussion focusses on moments or quantiles. For horizons of four centuries, there is a potential order of magnitude difference in valuation when $Q = 7$, and two orders of magnitude difference when $K = 4$. This illustrates the sensitivity of cost-benefit analysis to small differences in the choice of assumptions about the primitives of the discounting problem.

Panel D of Table 1 considers instead different families of distributions for $f(x)$. Compared to Panels B and C, at $H = 100$ years, valuations are relatively insensitive to the choice made. However, at $H = 200$ years, the valuation under the Wald distribution is less than half that of the baseline model, with the raw sample frequency distribution giving a valuation that is an order of magnitude larger. At 400 years, the differences are stark. Of particular note, as the discount rate for the sample frequency distribution is negative for such long time horizons, the \$1 million gets compounded under this model to a present value of \$76.8m. By contrast, the Wald distribution values the same cash flow at only two thousand dollars.

Table 2 is similar to Table 1, but now for the social cost of carbon (SCC) in US dollars per ton of carbon (\$/tC) calculated using each of the term structures reported in Figures 1 and 2, and based on the schedule of marginal carbon damages provided by Newell & Pizer (2003). These cash flows have a horizon of 400 years, with 50% of the undiscounted costs arising by year 170. This table also reports the present values calculated for an infrastructure project; the official estimates of benefits from Phase 1 of the HS2 rail link (London to Birmingham) that is currently being planned in the UK. The cash flows are

taken directly from the HS2 official website. These arise over a 75 year period to 2085, with 50% of the undiscounted benefits occurring by year 53. Finally, the table gives the present values of the estimated costs of decommissioning nineteen nuclear power stations in the UK as given in the Nuclear Decommissioning Authority (NDA) report and account 2012/13. While these span over a longer time horizon than HS2 (125 years), the ‘half-life’ is shorter, with 50% of the undiscounted costs occurring by year 29. Further details on these cash flow estimates are described in the technical online appendices of Freeman & Groom (2015).

[Insert Table 2 around here]

As the damages caused by greenhouse gas emissions are spread over such long time horizons, it is perhaps unsurprising that the estimated SCC is highly sensitive to the schedule of R_H that is deployed. For the greatest considered level of agreements over both moments and quantiles, $K = 4$ and $Q = 7$, the maximum value is three times the minimum value. When there is agreement over either two moments or three quantiles, the difference is more than ten-fold. Even when $f(x)$ is restricted to come from either a Wald or gamma distribution with parameters estimated through a method of moments approach, the latter valuation is more than 35% greater than the former. The sample frequency distribution gives the highest valuation at almost \$1,000/tC.

As a consequence of the shorter time periods covered by the HS2 rail link and nuclear decommissioning, the potential range of valuations are narrower in these cases than for the SCC. When consensus is on the quantiles, the bounds for R_H are widely spread even for low H ; see Figure 2. Given this, more precise valuations can be given for these examples when discussion is focussed around the moments instead. Agreement on the distributional family reduces the range of possible values even further, even when the sample frequency distribution is included as a viable characterization of $f(x)$.

3 The theoretical basis for DDRs

In the previous section, equation 1 was viewed through the lens of gamma discounting so as to make our central point as simply as possible. In this section we explain the theoretical underpinnings of this approach and some of the criticisms that have been raised in relation to it. We then show that other more conventional theoretical frameworks have the same structure as equation 1, with x interpreted differently in each case.

Gamma discounting stems from using the average expert discount factor to determine the certainty equivalent social discount rate. Yet the link between equation 1 and expected social welfare optimization is not immediately clear. As a consequence, it does not obviously follow that this approach is superior to using either the mean or median of x , the Expected Net Future Value criterion of Gollier (2009), or some other aggregation method or social choice criterion. In addition, recent evidence by Drupp et al. (2015) shows that such expert opinions on the SDR contain a mix of subjective normative judgements on matters of ethics, and positivist forecasts of verifiable quantities such as future interest rates or growth. As shown by Freeman & Groom (2015), and discussed below, the theory that lies behind aggregating these two types of opinion are very different. Nevertheless, theoretical justifications do exist for the aggregation of expert opinions using equation 1, although these are not entirely general and the assumptions are often restrictive.

3.1 Aggregating normative expert opinions

First consider the situation in which differences in expert opinion reflect heterogeneous values of the normative parameters of the Ramsey rule and associated SRTP: $x_i = \rho_i + \gamma_i g$. In this case, equation 1 generally emerges from frameworks where all experts have identical time separable and stationary logarithmic preferences ($\gamma_i = 1$ for all i) and beliefs about future consumption growth, g . Where experts differ is in their preferred pure rates of time preference, ρ_i .

In this context Jouini et al. (2010) and Jouini & Napp (2014) imagine heterogeneous agents each with individual endowments w_i . They assume that these agents are experts and show that if these experts were committed to trade inter-temporally within such an economy on the basis of these preferences and endowments, the equilibrium consumption path would be characterized by a social discount rate of the following form:

$$R_H^{JMN} = -\frac{1}{H} \ln \left[\sum_{i=1}^N \frac{w_i \rho_i}{\sum_{j=1}^n w_j \rho_j} \exp(-H x_i) \right] \quad (6)$$

This aggregation of agents preferences and beliefs has a similar form to equation 2, and equation 3 emerges when $w_i \rho_i = 1/N$ for all agents and the sample frequency of opinions is gamma distributed.

While equation 6 reflects the *equilibrium* social discount rate, Heal & Millner (2014) consider instead the social optimal in an economy with heterogeneous agents each with their own pure time preference, ρ_i , and Pareto weight w_i . A representative agent is sought that would choose the social optimal and in some sense reconcile the disagreement on ρ_i . Assuming again that each agent has logarithmic preferences ($\gamma_i = 1$ for all i), the social optimal can be implemented by a representative agent whose pure rate of time preference for aggregate utility for the time horizon H is given by:¹⁶

$$\rho_H^* = -\frac{1}{H} \ln \left[\sum_{i=1}^N w_i \exp(-\rho_i H) \right] \quad (7)$$

with associated social Ramsey rule, $x_H^* = \rho_H^* + g$. The price of a claim on \$1 at time H , p_H

¹⁶This equation differs from the one in Heal & Millner (2014) because they give instantaneous (forward) discount rates, while ρ_H^* here represents the (spot) discount rate between time zero and H . We use spot rates to be consistent with the rest of this paper. Further details of this distinction are available on request from the authors.

is therefore given by:

$$\begin{aligned}
p_H &= \exp(-Hx_H^*) = \exp(-H\rho_H^*) \exp(-Hg) \\
&= \sum_{i=1}^N w_i \exp(-\rho_i H) \exp(-Hg) \\
&= \sum_{i=1}^N w_i \exp(-x_i H)
\end{aligned} \tag{8}$$

This has the same functional form as equation 2 and yields equation 3 when $w_i = 1/N$ for all i .

3.2 Aggregating positivist expert opinion

Alternatively, the responses x_i might reflect different expert forecasts of a positive return to capital over the horizon of the cash flow. Define p_{tH} to be the price at time t of a claim on \$1 at time H (so that $p_{0H} = p_H$ and $p_{HH} = 1$), and x_{tH} as the per-period discount rate:

$$p_{t-1H} = E_{t-1} [p_{tH}] \exp(-x_{tH}) \tag{9}$$

Then by repeated iteration of this single-period present value equation (see, for example Ang & Liu (2004)):

$$\begin{aligned}
p_{0H} &= E_0 [p_{1H}] \exp(-x_{0H}) \\
&= E_0 [E_1 [p_{2H} \exp(-x_{1H}) \exp(-x_{0H})]] \\
&= \dots = E_0 [\exp(-H\bar{x}_H)]
\end{aligned} \tag{10}$$

where $\bar{x}_H = H^{-1} \sum_{t=0}^{H-1} x_{tH}$ is the average single-period expected rate of return to the claim over the horizon of the project. This has the same functional form as equation 1. For the purposes of this paper, we refer to equation 10 as the “Expected Net Present Value” condition, following Weitzman (1998). More detailed theoretical discussions of the derivation

of this approach have been given in Traeger (2013), Gollier & Weitzman (2010), Freeman (2010) and Gollier (2009).

Freeman & Groom (2015) consider the situation when the response of each expert, x_i is their own personal forecast of \bar{x}_H : $x_i = E_i[\bar{x}_H]$. In this case, the social planner can aggregate the individual expectations and produce a significantly more accurate assessment of the true value of \bar{x}_H than any of the individual experts. If individual forecast errors are identically and independently normally distributed, then it is the standard error of the distribution of responses that is the appropriate measure of remaining uncertainty, not the standard deviation. As a consequence, the social planner's pdf, $f(x)$, is much more heavily centered around the mean value than the sample frequency distribution of responses, effectively placing less weight on the individuals who have outlying opinions. This positivist interpretation of expert judgements therefore does not lead to the equal weighting of discount factors that underlies the gamma discounting approach that is represented by equation 3, and the resulting term structures of discount rates only decline slowly with the time horizon.

3.3 The econometric ENPV approach

Rather than seeking opinions, the pdf of \bar{x}_H in equation 10 can instead be estimated using an econometric approach. In order to do this it is necessary to associate x_{tH} with the return on a specific financial asset. The standard choice is to set $x_{tH} = r_{ft}$, the yield on a risk-free Treasury bond. Yet this choice is, in itself, controversial and only follows under certain restrictive assumptions. This is because, even though the final payoff is known to be \$1, the interim price of the claim on this cash flow, p_{tH} for $t \in [1, H - 1]$ is unknown in the previous period. Therefore the one period discount rate, x_{tH} , will, in general, incorporate a risk premium reflecting this uncertainty over p_{tH} . Applying a single period risk-free rate is only appropriate under conditions that endogenize a zero risk premium. Most notably Cox et al. (1981) argue that “locally certainty” in consumption (c_{t+1} being perfectly known at time t) justifies this choice. Gollier (2014), in particular, has argued that this assumption

is highly unrealistic.

Despite its theoretical limitations, the econometric ENPV approach has been employed in a number of highly-cited studies on long-term discount rates; see, for example, Newell & Pizer (2003), Groom et al. (2007), Gollier et al. (2008), Hepburn et al. (2009), and Freeman et al. (2015). These, in turn, have been highly influential in shaping international governmental policy. We return to this framework in the next section.

3.4 Consumption-based asset pricing

Equation 1 can alternatively be interpreted through consumption-based approaches to social discount rates. Suppose utility $u(c_t, t)$ is gained from consuming c_t units of the single consumption good at time t . From a standard Euler equation, if a project makes a future payment of \$1 at time H and nothing at any other time, then its present value is given by:

$$p_H = \frac{E[u'(c_H, H)]}{E[u'(c_0, 0)]} \quad (11)$$

Proposition 1 describes how equation 1 can be generalized to the consumption-based approach.

Proposition 1: Assuming current consumption, c_0 , is non-stochastic and that there is a time-separable power utility function of the form, $u'(c_H, H) = e^{-\rho H} c_H^{-\gamma}$, with pure time preference rate ρ and coefficient of relative risk aversion γ , then:

$$\begin{aligned} p_H &= E[\exp(-Hx_H)] \\ x_H &= \rho + \frac{\gamma}{H} \ln(c_H/c_0) \end{aligned} \quad (12)$$

and this also takes the same functional form as equation 1.

Proof: See Appendix B

In summary, there are a number of different theoretical justifications for equation 1, and while each requires that x is interpreted in different ways, the essential mathematical structure remains the same. Showing that this structure has a range of theoretical interpretations gives our results broader relevance and the potential for more applications. Much, but not all, of the remaining theoretical contention reflects either (i) the choice to weight all discount factors equally when deriving equation 3, and this explains why these weights were the focus of our attention in the previous section, or (ii) associating x_{tH} with a yield on a risk-free Treasury bond.

4 Other sharp bounds term structures

In this section, we take the econometric ENPV and consumption-based asset pricing approaches to equation 1, which have both been influential in policy circles, and create sharp bounds for the certainty-equivalent social discount rate in each case.

4.1 The econometric ENPV approach

We follow the state-space model of Groom et al. (2007) within the econometric ENPV approach. Let $\theta_{ft} = \ln(100r_{ft})$, and assume that this evolves according to:

$$\theta_{ft} = \eta + \lambda_t \theta_{ft-1} + e_t \tag{13}$$

$$\lambda_t = \eta_1 \lambda_{t-1} + u_t$$

which is an AR(1) process with time-varying autoregressive parameter, λ_t . The error terms are independently and identically normally distributed (i.i.n.d.) with variance σ_e^2 and σ_u^2 respectively and zero means. Parameter calibrations are taken directly from Groom et al. (2007), and the properties of $f_H(x_H)$, which are now H -dependent, are then determined by

simulation.¹⁷

The top graph in Figure 3 presents values of R_H under this econometric ENPV interpretation of equation 1, with the solid line being directly comparable against Figure 3 in Groom et al. (2007). The graph also presents sharp upper and lower bounds for R_H when there is agreement on the upper and lower supports of $f_H(x_H)$ as well as the first K moments, once more determined using the technique in Karlin & Studden (1966). Again it is clear that, for long time horizons, R_H can potentially lie within a wide range, although the bounds are not as spread as in Figure 1.

[Insert Figure 3 around here]

Panel A of Table 3 mimics Panels A–C of Table 2 when x_H is interpreted within the framework of Groom et al. (2007). It reports the SCC and the present values of HS2 and nuclear decommissioning under the baseline model, and the sharp upper and lower bounds of R_H for different moments (K) and quantiles (Q).

[Insert Table 3 around here]

The ranges of possible values are narrower now than in Table 2 because the upper and lower bounds for R_H are tighter in Figure 3 than Figure 1. For decision making purposes, there is still considerable uncertainty over the SCC; ranging from \$16.2/tC to \$22.6/tC when $K = 3$ and \$14.8/tC to \$20.4/tC when $Q = 7$.

4.2 Consumption-based asset pricing

To derive declining discount rates in the consumption-based asset pricing setting, suppose $\ln(c_t/c_{t-1}) = \mu + \varepsilon_t$, where $\varepsilon_t \sim N(0, \sigma^2)$, but that the value of μ is unknown. This

¹⁷The parameter estimates (with associated standard errors) are $r_{f0} = 4\%$, $\eta = 0.510$ (0.0082), $\eta_1 = 0.990$ (0.002), $\ln(\sigma_e^2) = -9.158$ (1.324), $\ln(\sigma_u^2) = -6.730$ (0.144). Further details of the simulation process are available on request from the authors. Unreported results are also available for a parameterization of the econometric ENPV model that is based on a calibration reported in Newell & Pizer (2003). Again, our central conclusion holds in this case.

approach is used as a justification for the French government’s position on long-term social discounting, although the arguments are based on a simple numerical example rather than an empirical analysis; see (Lebegue 2005, p.102).¹⁸ To calibrate this model we follow Gollier (2012), and assume μ takes one of two values, $\mu_u = 3\%$ or $\mu_l = 1\%$ with equal probability, making $f_H(x_H)$ a mixture normal distribution with properties that depend on H .

In the bottom graph of Figure 3, we again present R_H under the baseline model, which can be compared against Figure 6.2 of Gollier (2012), as well as upper and lower bounds of R_H contingent on agreement over the support and first K moments of $f_H(x_H)$.¹⁹ When $K = 4$, the bounds in this case are quite tight at all horizons, but are then much more widely spread for lower values of K .

Panel B of Table 3 gives the present values under this model. Generally, except for $K = 4$, the potential range of values is more dispersed than under the econometric ENPV model reported in Panel A, but narrower than for the gamma discounting calibrations reported in Table 2. Strong agreement on the moments continues to give a more precise valuation than when consensus is over the quantiles. But in all cases, it is not possible to precisely determine the SCC.

5 Conclusion

There are a number of theoretical and empirical arguments for using a declining term structure of social discount rates for intergenerational projects. These largely stem from incorporating uncertainty in future consumption growth or interest rates into the analysis, or dealing with disagreement between experts on the discount rate. So persuasive have these

¹⁸It is well known that DDRs require persistence in the growth diffusion process in this setting. If, alternatively, growth is independently and identically distributed the term structure will be flat. See, for example, (Gollier 2012, Ch 3).

¹⁹This graph is based on parameter values $\rho = 0$, $\gamma = 2$ and $\sigma = 3.6\%$. Further details of the calculations are available on request for the authors. Unreported results are also available for the sharp upper and lower bounds of R_H in the Markov regime switching models which underlie Figures 5.2 and 5.3 in Gollier (2012) in this consumption based asset pricing setting framework. Again, the main conclusions of this paper hold within this setting.

arguments been that they are now recognized in government policies and recommendations in the UK, France, Norway and Denmark. The practical question that necessarily follows is: how can the theory be operationalized?

So far policy makers have sought a range of expert economists' advice on the best route forward, and the experts have provided them with empirical estimates of the certainty-equivalent term structure of social discount rates based on specific theories and characterizations of the future (e.g. Gollier (2012), Arrow et al. (2014)). The fact of the matter is, though, that we know far less about what lies ahead than the complete characterizations of risk that many of these models suggest. This raises the question, how certain are we about the certainty-equivalent social discount rate?

This paper shows that, even if there are strongly overlapping views on the primitives of social discounting, the empirical term structure of the certainty-equivalent social discount rate emerging from many theoretical models cannot be positioned within anything but very wide bounds. This is particularly true for long time horizons. The obvious implication of this is that policy prescriptions become less crisp for all but the highest (or lowest) return public projects. For instance, viable estimates for the social cost of carbon and the net present values of other intergenerational project become alarmingly dispersed from this position of ignorance. Apparently trivial disagreements over parameterization choices can lead to significant differences in policy recommendations.

While not quite a dismal theorem, overall this paper presents a depressing finding for practitioners of cost-benefit analysis. Even if we are willing to put issues of Knightian uncertainty to one side, we must accept that we know little about the true nature of what the distant future might hold. This admission of ignorance means that estimated present values are likely to be so imprecise as to provide only minimal guidance to policy makers on intergenerational projects. One positive outcome of the paper is the formal mechanism for determining the range over which sensitivity analysis might be undertaken when doubt exists over the appropriate choice of discount rate. However, the broader message is that

we may have to look towards additional decision-making criteria, beyond the mechanical calculation of Net Present Values, to shed more light on the social value of intergenerational projects.

Appendix A: Generating the sharp bounds

In this appendix, we describe in more detail the method that we use for determining the upper and lower bounds of R_H when there is agreement on the supports and moments of $f_H(x_H)$. The H -subscripts in this appendix reflect the fact that x_H may be horizon dependent in some interpretations of equation 1, although not gamma discounting.

Consider the set of all well-defined probability density functions, \mathfrak{S}_H , with elements g_H , which are supported on a common interval $[a_H, b_H]$.²⁰ We assume that there is consensus that the ‘true’ $f_H(x_H)$ is an element of this set, but we do not know which element it is.

Next we suppose that we only know the first K (non-central) moments of $f_H(x_H)$; $E_f[x_H^k] = M_{kH}$ for $k \leq K$ where $E_f[\cdot]$ is the expectation operator conditional on the pdf of x_H being f_H . The smaller is K , the more ignorant we are about $f_H(x_H)$, but claiming any knowledge of the moments of the pdf allows us to narrow our search to the subset $\mathfrak{S}_{KH} \subset \mathfrak{S}_H$ that contains all elements g_H with first K moments $E_g[x_H^k] = M_{kH}$ for $k \leq K$. We then define strict upper and lower bounds for R_H , R_{uH} and R_{lH} , by:

$$\begin{aligned} R_{uH} &= -\frac{1}{H} \ln(\inf[E_g[\exp(-Hx_H)] | g_H \in \mathfrak{S}_{KH}]) \\ R_{lH} &= -\frac{1}{H} \ln(\sup[E_g[\exp(-Hx_H)] | g_H \in \mathfrak{S}_{KH}]) \end{aligned} \tag{14}$$

As expressions of the form $E[\exp(-Hx_H)]$ are moment generating functions (mgf), or Laplace-Stieltjes transformations, we can invoke a powerful result from Karlin & Studden (1966) to find R_{uH} and R_{lH} . These are derived by establishing two separate, discrete pdfs

²⁰There is a restriction that a_H is finite. For all the examples we consider we also take finite b_H , but the extension to infinite b_H is straightforward given the results in Eckberg (1977).

which are, loosely speaking, ‘at opposite ends’ of the support and yet share the first K moments. The upper bound for the mgf is found by calculating the most extreme discrete distribution to place as much mass as possible in the left hand tail (lower values of the discount rate), while still satisfying the K moment conditions. The lower bound is found by minimizing the mass in the left hand tail.

More concretely, the extreme discrete distributions place non-zero probability mass at ϖ points on the interval $[a_H, b_H]$ where the number of mass points depends on the number of moments, K , of the distribution of x_H that we are willing to assume we agree upon: $\varpi \in [(K+1)/2, (K+3)/2]$. We denote these points by V_{qlH} (v_{qlH}) for the lower bound of the mgf and V_{quH} (v_{quH}) for the upper bound when $a_H = 0$ ($a_H \neq 0$), with associated probabilities π_{qlH} and π_{quH} , where q indexes the mass points from smaller to larger values of x_H .

A.1 $a_H = 0$

First, consider the restricted case when $a_H = 0$ and $b_H = B_H$; we broaden the discussion to more general values of a_H and b_H in the next subsection. The method of Karlin & Studden (1966) now follows from the observation that the set of functions $\left\{1, x_H, \dots, x_H^K, (-1)^{K+1} \exp(-Hx_H)\right\}$ for $H > 0$ and positive integer K is a Tchebycheff system. This allows us to identify the properties of the discrete distributions that give sharp bounds for the mgf (see, for example, Eckberg (1977)):

	V_{qlH}	V_{quH}
K even	$\varpi = (K + 2) / 2$ $K/2$ points in $(0, B_H)$ One point at B_H	$\varpi = (K + 2) / 2$ $K/2$ points in $(0, B_H)$ One point at 0
K odd	$\varpi = (K + 1) / 2$ $(K + 1) / 2$ points in $(0, B_H)$	$\varpi = (K + 3) / 2$ $(K - 1) / 2$ points in $(0, B_H)$ One point at 0 One point at B_H

Each extreme discrete distribution is uniquely identified since the number of degrees of freedom equals the number of moment constraints. For example, for even K , there are $K/2$ degrees of freedom on location and $(K + 2) / 2 - 1$ degrees of freedom for the probabilities, giving a total number of degrees of freedom of K . It is also straightforward to verify that there are K degrees of freedom in parameter choice when K is odd. Therefore the extreme density functions are uniquely defined by the K moment conditions in all cases.

Closed form solutions for V_{qlH} , V_{quH} , π_{qlH} and π_{quH} are available for $K \leq 3$. When $K = 1$, the lower bound has only one mass point which is on the mean value; $V_{1lH} = M_{1H}$. The upper bound has mass at $V_{1uH} = 0$ and $V_{2uH} = B_H$ only and the probability π_{1uH} is set to ensure that the mean is equal to M_{1H} ; $\pi_{1uH} = (B_H - M_{1H}) / B_H$. Closed form solutions for $K = 2$ and $K = 3$ are given in Eckberg (1977). For the case $K = 2$:

	$q = 1$	$q = 2$		$q = 1$	$q = 2$
V_{qlH}	$\frac{M_{1H}B_H - M_{2H}}{B_H - M_{1H}}$	B_H	V_{quH}	0	$\frac{M_{2H}}{M_{1H}}$
π_{qlH}	π_1	$1 - \pi_1$	π_{quH}	$\frac{M_{2H} - M_{1H}^2}{M_{2H}}$	$\frac{M_{1H}^2}{M_{2H}}$

where $\pi_1 = (B_H - M_{1H})^2 / (M_{2H} - M_{1H}^2 + (B_H - M_{1H})^2)$. For $K = 3$:

	$q = 1$	$q = 2$		$q = 1$	$q = 2$	$q = 3$
V_{qlH}	$\frac{A_1 - \chi}{2}$	$\frac{A_1 + \chi}{2}$	V_{quH}	0	ζ	B_H
π_{qlH}	$\frac{\chi + A_1 - 2M_{1H}}{2\chi}$	$\frac{\chi - A_1 + 2M_{1H}}{2\chi}$	π_{quH}	$1 - \pi_1^* - \pi_2^*$	π_1^*	π_2^*

where:

$$A_1 = \frac{M_{3H} - M_{1H}M_{2H}}{M_{2H} - M_{1H}^2}, \quad A_2 = \frac{M_{2H}^2 - M_{1H}M_{3H}}{M_{2H} - M_{1H}^2}, \quad \chi = \sqrt{A_1^2 + 4A_2}$$

$$\pi_2^* = \frac{M_{1H}M_{3H} - M_{2H}^2}{M_{1H}B_H^3 - 2M_{2H}B_H^2 + M_{3H}B_H}, \quad \pi_1^* = \frac{(M_{1H} - \pi_2^*B_H)^2}{M_{2H} - \pi_2^*B_H^2}, \quad \zeta = \frac{M_{2H} - \pi_2^*B_H^2}{M_{1H} - \pi_2^*B_H}$$

As noted by (Johnson & Taaffe 1993, p.96), “less analytically tractable cases (e.g., four or five non-central moments) call for use of symbolic or numerical methods for solving the nonlinear equations”. We use numerical methods here.²¹

A.2 $a_H \neq 0$

When $f_H(x_H)$ is supported on the interval $[a_H, b_H]$ for $a_H \neq 0$, we first undertake a change of variables. Define $y_H = x_H - a_H$, $g_H^*(y_H) = g_H(x_H)$ and \mathfrak{S}_{KH}^* as the set containing all elements $g_H^*(y_H)$. The lower bounds for R_H are derived from the bounds for $E_{g^*}[\exp(-Hy_H)]$.

$$R_{lH} = -\frac{1}{H} \ln(\exp(-Ha_H) \sup[E_{g^*}[\exp(-Hy_H)] | g_H^* \in \mathfrak{S}_{KH}^*])$$

and there is an analogous expression for R_{uH} , with the supremum replaced by the infimum.

²¹Details available on request from the authors.

Appendix B: Proof of Proposition 1

Equation 11 states that the price, p_H , of a payoff of \$1 at time H is given by:

$$p_H = \frac{E[u'(c_H, H)]}{u'(c_0, 0)}$$

assuming c_0 is non-stochastic. With the utility function defined by $u'(c_H, H) = e^{-\rho H} c_H^{-\gamma}$:

$$p_H = E \left[e^{-\rho H} \left(\frac{c_H}{c_0} \right)^{-\gamma} \right]$$

Using offsetting exponential and logarithm functions, the fact that $\gamma \ln(z) = \ln(z^\gamma)$, and then factoring out H yields:

$$\begin{aligned} p_H &= E \left[e^{-\rho H} \exp \left\{ -\gamma \ln \left(\frac{c_H}{c_0} \right) \right\} \right] \\ &= E \left[\exp \left\{ -\rho H - \gamma \ln \left(\frac{c_H}{c_0} \right) \right\} \right] \\ &= E \left[\exp \left\{ -H \left(\rho + \frac{\gamma}{H} \ln \left(\frac{c_H}{c_0} \right) \right) \right\} \right] \end{aligned} \tag{15}$$

Then p_H can be written as $p_H = E[\exp(-Hx_H)]$ where $x_H = \rho + \frac{\gamma}{H} \ln \left(\frac{c_H}{c_0} \right)$, which can be interpreted as the Social Rate of Time Preference for an annualized growth rate given by $H^{-1} \ln \left(\frac{c_H}{c_0} \right)$. This is equation 12. QED.

References

- Almond, D. (2006), ‘Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population’, *Journal of Political Economy* **114**, 672–712.
- Ang, A. & Liu, J. (2004), ‘How to discount cashflows with time-varying expected returns’,

- Journal of Finance* **59**, 2745–2783.
- Arrow, K., Cropper, M., Gollier, C., Groom, B., Heal, G., Newell, R., Nordhaus, W., Pindyck, R., Pizer, W., Portney, P., Sterner, T., Tol, R. & Weitzman, M. (2013), ‘Determining benefits and costs for future generations’, *Science* **341**, 349–350.
- Arrow, K., Cropper, M., Gollier, C., Groom, B., Heal, G., Newell, R., Nordhaus, W., Pindyck, R., Pizer, W., Portney, P., Sterner, T., Tol, R. & Weitzman, M. (2014), ‘Should governments use a declining discount rate in project analysis?’, *Review of Environmental Economics and Policy* **8**, 145–163.
- Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1981), ‘A re-examination of traditional hypotheses about the term structure of interest rates’, *Journal of Finance* **36**, 769–799.
- Cropper, M., Freeman, M., Groom, B. & Pizer, W. (2014), ‘Declining discount rates’, *American Economic Review: Papers & Proceedings* **104**, 538–543.
- Dasgupta, P. (2008), ‘Discounting climate change’, *Journal of Risk and Uncertainty* **37**, 141–169.
- Drupp, M. A., Freeman, M. C., Groom, B. & Nesje, F. (2015), Discounting disentangled: An expert survey on the determinants of the long-term social discount rate. Grantham Research Institute Working Paper No. 172, London School of Economics.
- Eckberg, A. E. (1977), ‘Sharp bounds on Laplace-Stieltjes transforms, with applications to various queueing problems’, *Mathematics of Operations Research* **2**, 135–142.
- Freeman, M. C. (2010), ‘Yes, we should discount the far-distant future at its lowest possible rate: A resolution of the Weitzman Gollier puzzle’, *Economics: The Open-Access, Open-Assessment E-Journal* **4**(2010–13).
- Freeman, M. C. & Groom, B. (2013), ‘Biodiversity valuation and the discount rate problem’, *Accounting, Auditing and Accountability Journal* **26**, 715–745.

- Freeman, M. C. & Groom, B. (2015), ‘Positively gamma discounting: Combining the opinions of experts on the social discount rate’, *Economic Journal* **125**, 1015–1024.
- Freeman, M. C., Groom, B., Panopoulou, E. & Pantelidis, T. (2015), ‘Declining discount rates and the Fisher Effect: Inflated past, discounted future?’, *Journal of Environmental Economics and Management* **73**, 32–49.
- Gollier, C. (2009), ‘Should we discount the far-distant future at its lowest possible rate?’, *Economics - The Open-Access, Open-Assessment E-Journal* **3**(2009–25), 1–14.
- Gollier, C. (2012), *Pricing the Planet’s Future: The Economics of Discounting in an Uncertain World*, Princeton University Press, Princeton.
- Gollier, C. (2014), Gamma discounters are short-termist. IDEI Working Paper, no. 828.
- Gollier, C., Koundouri, P. & Pantelidis, T. (2008), ‘Declining discount rates: Economic justifications and implications for long-run policy’, *Economic Policy* **23**(56), 757–795.
- Gollier, C. & Weitzman, M. L. (2010), ‘How should the distant future be discounted when discount rates are uncertain?’, *Economics Letters* **107**(3), 350–353.
- Gollier, C. & Zeckhauser, R. (2005), ‘Aggregation of heterogeneous time preferences’, *Journal of Political Economy* **113**, 878–896.
- Groom, B., Koundouri, P., Panopoulou, E. & Pantelidis, T. (2007), ‘Discounting the distant future: How much does model selection affect the certainty equivalent rate?’, *Journal of Applied Econometrics* **22**(3), 641–656.
- Groom, B. & Maddison, D. J. (2013), Non-identical quadruplets: Four new estimates of the elasticity of marginal utility for the UK. Grantham Institute Centre for Climate Change Economics and Policy Working Paper No. 141.
- Heal, G. & Millner, A. (2014), ‘Agreeing to disagree on climate policy’, *PNAS* **111**, 3695–3698.

- Hepburn, C., Koundouri, P., Panopoulou, E. & Pantelidis, T. (2009), ‘Social discounting under uncertainty: A cross-country comparison’, *Journal of Environmental Economics and Management* **57**(2), 140–150.
- Iverson, T. (2013), ‘Minimax regret discounting’, *Journal of Environmental Economics and Management* **66**, 598–608.
- Johnson, M. A. & Taaffe, M. R. (1993), ‘Tchebycheff systems for probabilistic analysis’, *American Journal of Mathematical and Management Sciences* **13**, 83–111.
- Jouini, E., Marin, J.-M. & Napp, C. (2010), ‘Discounting and divergence of opinion’, *Journal of Economic Theory* **145**, 830–859.
- Jouini, E. & Napp, C. (2014), ‘How to aggregate experts’ discount rates: An equilibrium approach’, *Economic Modelling* **36**, 235–243.
- Karlin, S. & Studden, W. (1966), *Tchebycheff Systems: With Applications in Analysis and Statistics*, John Wiley & Sons, New York.
- Klibanoff, P., Marinacci, M. & Mukerji, S. (2011), ‘Definitions of ambiguous events and the smooth ambiguity model’, *Economic Theory* **48**, 399–424.
- Knight, F. H. (1921), *Risk, Uncertainty and Profit*, Hart, Schaffner & Marx; Houghton Mifflin Company, Boston MA.
- Lebegue, D. (2005), Revision du taux de actualisation des investissements publics. rapport du groupe de experts, commissariat generale de plan., Technical report, Commissariat Generale de Plan.
- Millner, A. & Heal, G. (2014), Resolving intertemporal conflicts: Economics vs. politics. NBER Working Paper No. 20705.

- Newell, R. G. & Pizer, W. A. (2003), ‘Discounting the distant future: How much do uncertain rates increase valuations?’, *Journal of Environmental Economics and Management* **46**, 52–71.
- Nordhaus, W. D. (2007), ‘A review of the Stern Review on the Economics of Climate Change’, *Journal of Economic Literature* **45**, 686–702.
- O’Hagan, A. (2012), ‘Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux’, *Environmental Modelling and Software* **36**, 35–48.
- O’Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E. & Rakow, T. (2006), *Uncertain Judgements: Eliciting Experts’ Probabilities*, Statistics in Practice, John Wiley and Sons, Chichester, England.
- Pindyck, R. S. (2015), The use and misuse of models for climate policy, Working Paper 21097, National Bureau of Economic Research.
- Stern, N. (2008), ‘The economics of climate change’, *American Economic Review: Papers & Proceedings* **98**, 1–37.
- Traeger, C. P. (2013), ‘Discounting under uncertainty: Disentangling the Weitzman and Gollier effect’, *Journal of Environmental Economics and Management* **66**, 573–582.
- Traeger, C. P. (2014), ‘Why uncertainty matters — discounting under intertemporal risk aversion and ambiguity’, *Economic Theory* **56**, 627–664.
- Weitzman, M. L. (1998), ‘Why the far-distant future should be discounted at its lowest possible rate’, *Journal of Environmental Economics and Management* **36**, 201–208.
- Weitzman, M. L. (2001), ‘Gamma discounting’, *American Economic Review* **91**, 260–271.
- Weitzman, M. L. (2009), ‘On modeling and interpreting the economics of catastrophic climate change’, *The Review of Economics and Statistics* **91**(1), 1–19.

Weitzman, M. L. (2011), ‘Fat-tailed uncertainty in the economics of catastrophic climate change’, *Review of Environmental Economics and Policy* **5**(2), 275–292.

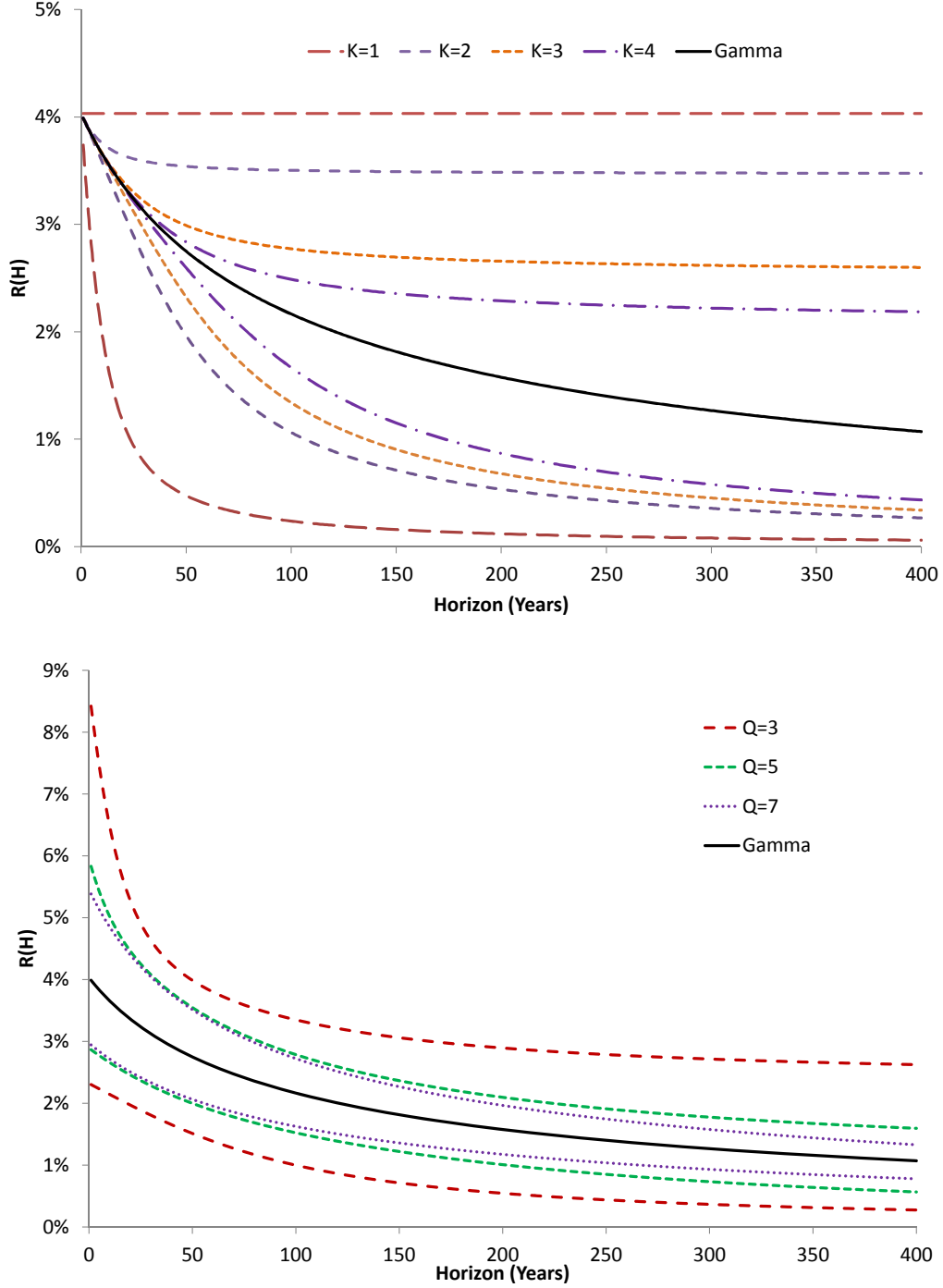


Figure 1. This figure presents the term structure of social discount rates as given through the gamma discounting schedule of Weitzman (2001). The solid line is the baseline parameterization from this model. We then present upper and lower bounds for R_H conditional on matching the first K moments of the distribution for $K \in [1, 4]$ in the top graph, and Q quantiles of the distribution, for $Q \in \{3, 5, 7\}$ in the bottom graph. The support of the probability density function in each case is $[0\%, 19.11\%]$.

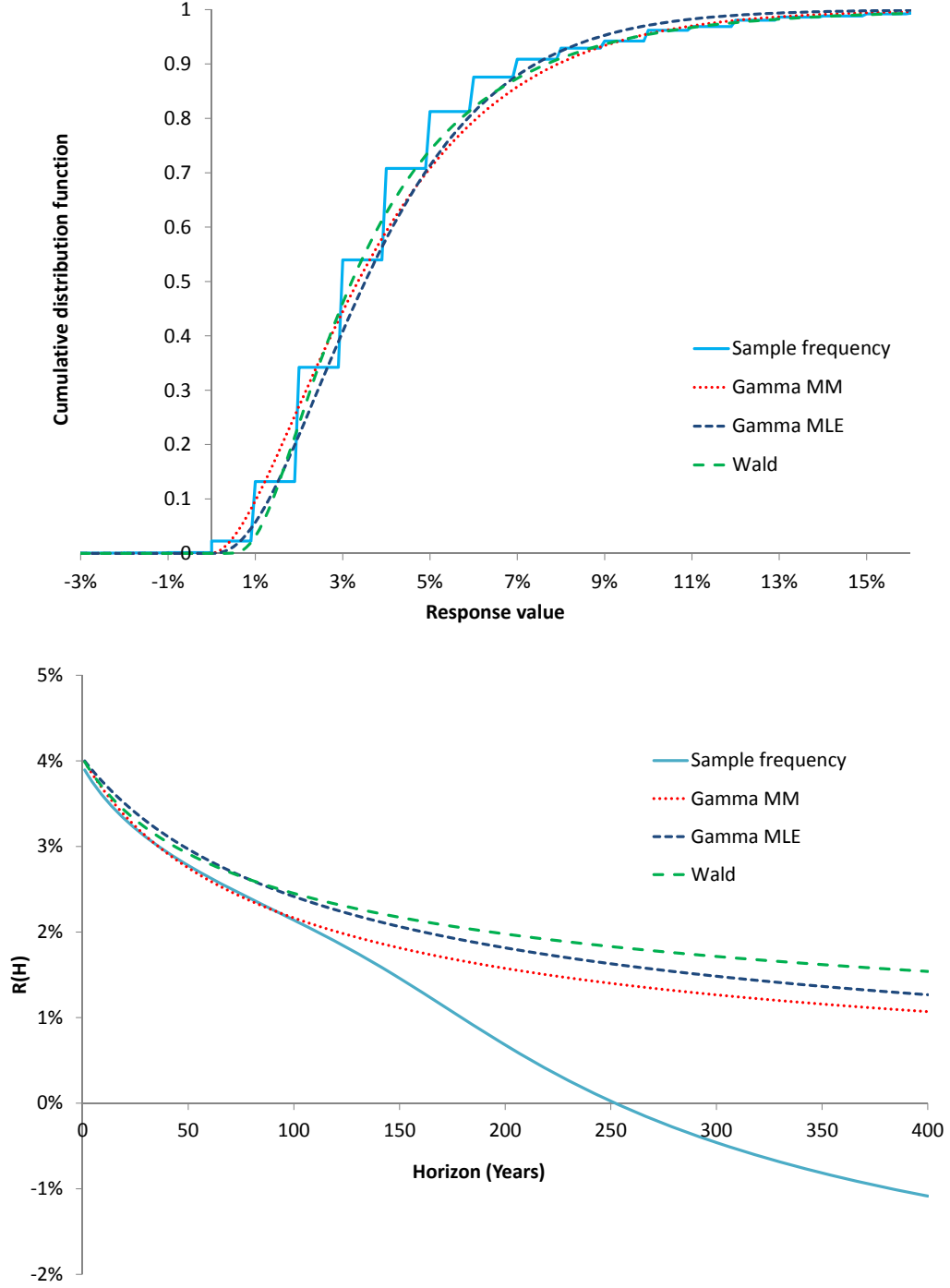


Figure 2. The top graph presents the cumulative distribution function (cdf) for the raw sample frequency distribution of Weitzman’s gamma discounting survey data. It also shows cdfs of gamma distributions fitted to this data using method of moments (MM) and maximum likelihood (MLE) techniques, as well as the cdf of a Wald (Inverse Gaussian) distribution fitted by a method of moments approach to the same data. The bottom graph gives the schedule of discount rates for each of these distributions.

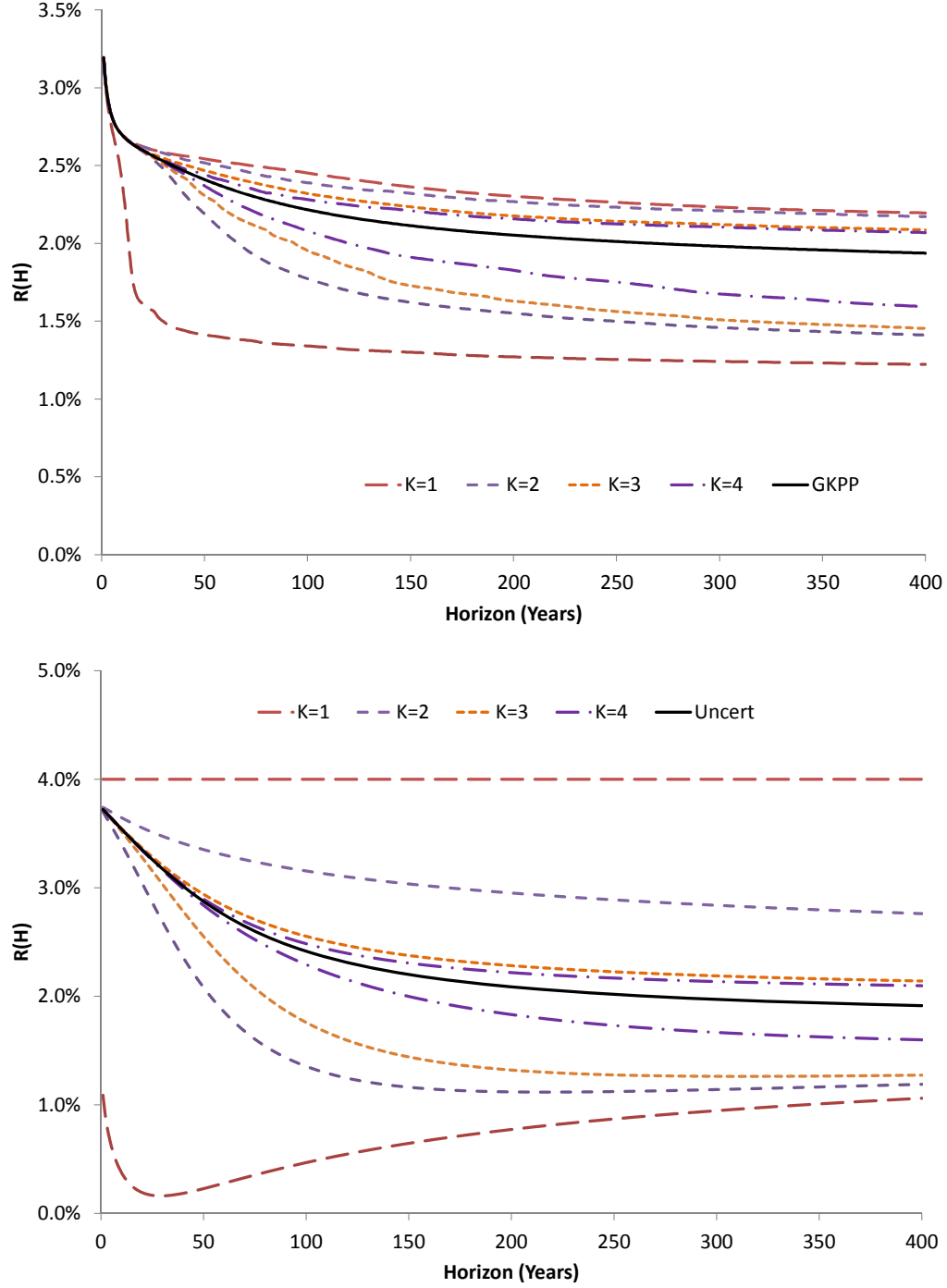


Figure 3. The top graph presents the term structure of social discount rates as given through the ENPV setting by the state-space model of Groom et al. (2007), while the bottom graph is derived from the parameter uncertainty model for logarithmic consumption growth described by Gollier (2012). The solid lines are the baseline parameterization from the models, which can be compared against Figure 3 in Groom et al. (2007) and Figure 6.2 of Gollier (2012). The graphs also present upper and lower bounds for R_H conditional on matching the first K moments of the underlying distribution.

	$H = 100$ years		$H = 200$ years		$H = 400$ years	
	Min	Max	Min	Max	Min	Max
Panel A: Gamma Method of Moments						
Baseline Model	114,778		42,782		13,839	
Panel B: Moments						
K=1	17,752	789,125	315	789,125	0.10	789,125
K=2	30,128	345,754	942	344,356	0.92	344,353
K=3	62,556	261,809	4,930	257,593	30.6	257,569
K=4	83,094	188,953	10,311	176,622	159.4	176,410
Panel C: Quantiles						
Q=3	35,117	368,450	3,067	336,400	27.6	333,361
Q=5	61,643	218,117	15,104	133,267	1,696	103,939
Q=7	65,672	196,761	19,627	95,628	4,913	44,402
Panel D: Distributional Family						
Gamma MLE	89,360		26,489		6,290	
Wald	86,130		19,168		2,114	
Sample frequency	117,736		255,955		76,778,160	

Table 1. The present value of \$1m at horizons of 100, 200 and 400 years under different discounting schedules.

	SCC		HS2		NDA	
	Min	Max	Min	Max	Min	Max
Panel A: Gamma Method of Moments						
Baseline Model	20.0		23.1		46.8	
Panel B: Moments						
K=1	5.3	188.9	13.8	66.7	38.9	85.2
K=2	7.2	83.7	16.7	33.2	41.7	55.2
K=3	11.1	63.8	20.7	28.3	44.8	51.2
K=4	13.6	46.1	22.1	25.0	45.9	48.6
Panel C: Quantiles						
Q=3	6.7	86.5	13.0	41.1	33.4	64.6
Q=5	10.8	43.1	16.1	32.6	38.5	56.4
Q=7	11.7	35.2	16.4	31.6	39.1	55.5
Panel D: Distributional Family						
Gamma MLE	15.6		21.0		45.0	
Wald	14.6		21.5		45.4	
Sample frequency	979.8		22.9		46.9	

Table 2. The present values, as calculated through the different discounting schedules presented in Figures 1 and 2, of (i) the Social Cost of Carbon (SCC) in terms of dollars per ton of carbon (\$/tC), (ii) the costs of Phase 1 of the High Speed 2 (HS2) rail line in the UK; London to Birmingham (£bn), (iii) the costs of decommissioning the previous generation of nuclear power stations in the UK (£bn).

	SCC		HS2		NDA	
	Min	Max	Min	Max	Min	Max
Panel A: ENPV						
K=1	14.8	39.8	25.5	42.3	50.7	64.8
K=2	15.4	25.8	25.8	29.9	51.0	54.1
K=3	16.2	22.6	26.3	28.3	51.4	52.9
K=4	16.6	19.9	26.7	27.6	51.6	52.3
Q=3	11.9	25.9	18.6	33.3	41.4	57.8
Q=5	14.6	21.5	23.3	30.3	47.7	55.0
Q=7	14.8	20.4	24.1	29.7	49.1	54.4
Model	17.5		27.0		51.9	
Panel B: Parameter Uncertainty						
K=1	5.4	89.2	13.9	74.1	39.1	94.1
K=2	8.6	38.9	18.0	31.3	42.9	53.8
K=3	12.9	27.3	21.3	25.4	45.5	49.0
K=4	13.6	16.5	21.7	22.3	45.9	46.4
Q=3	6.5	65.6	12.2	60.3	33.2	85.5
Q=5	8.7	32.5	15.0	37.2	37.6	62.2
Q=7	9.2	24.0	15.5	31.9	38.3	57.1
Model	14.4		21.9		46.1	

Table 3. As Table 2, but with bounds now based on the ENPV model of Groom et al. (2007) and the parameter uncertainty model for logarithmic consumption growth described by Gollier (2012). The terms structures used are presented in Figure 3.